Waveform Consistency Test in Burst Detection

Laura Cadonati

LSC meeting, Hannover

August 20, 2003

LIGO-G030438-00-Z

Test goal

The LIGO Burst Search pipeline uses *Event Trigger Generators* (ETGs) to flag times when “something anomalous” occurs in the strain time series

$\Rightarrow$ burst candidate events (Δt, Δf, SNR)

Events from the three LIGO interferometers are brought together in coincidence (time, frequency, power).

In order to use the full power of a coincident analysis:

» Are the waveforms consistent? To what confidence?

» Can we suppress the false rate in order to lower thresholds and dig deeper into the noise?

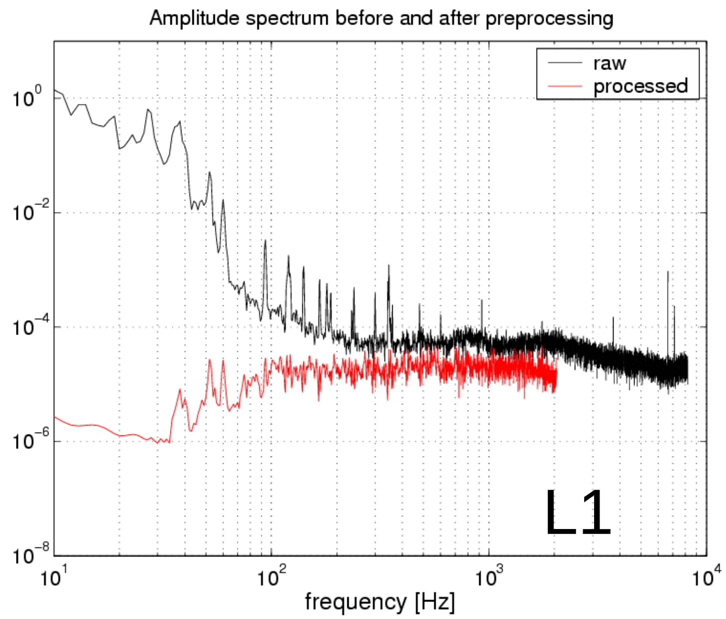
Cross correlation of coincident events

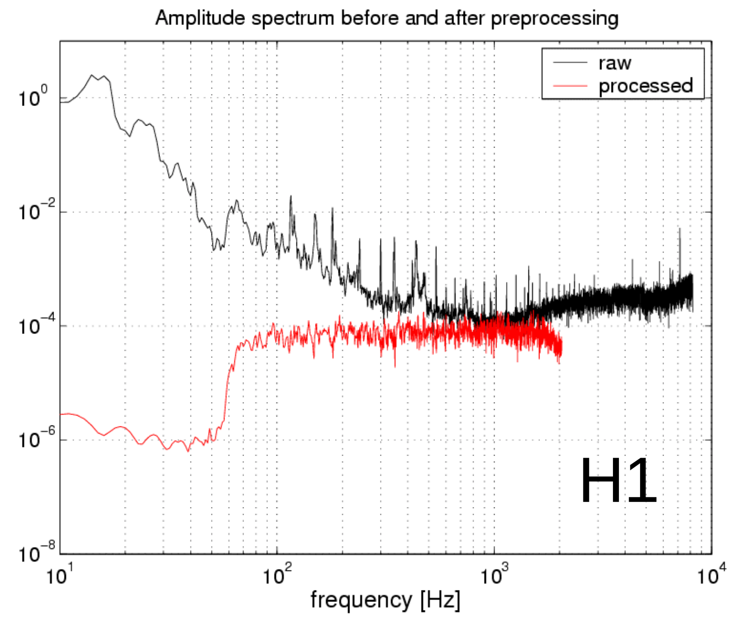
Data Conditioning

Decimate and high-pass few seconds of data around event ⇒100-2048 Hz

Remove predictable content (effective whitening/line removal): train a linear predictor error filter over 1 s of data (1 s before event start), apply as second order sections model using zero-phase filtering - described in S. Chatterji's talk on data conditioning (LIGO-G030439)

⇒emphasis on transients, avoid non-stationary, correlated lines.





r-statistic

$$r = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2} \sqrt{\sum_i (y_i - \bar{y})^2}}$$

Linear correlation coefficient or normalized cross correlation for the two series $\{x_i\}$ and $\{y_i\}$

NULL HYPOTHESIS: the two (finite) series $\{x_i\}$ and $\{y_i\}$ are uncorrelated

$\Rightarrow$ Their linear correlation coefficient (**Pearson's r**) is normally distributed around zero, with $\sigma = 1/\text{sqrt}(N)$ where N is the number of points in the series (N >> 1)

$$S = erfc\ (|r|\ \text{sqrt}(N/2))$$

double-sided significance of the null hypothesis

i.e.: probability that |r| is larger than what measured, if $\{x_i\}$ and $\{y_i\}$ are uncorrelated

$$C = -\log_{10}(S)$$

confidence that the null hypothesis is FALSE $\Rightarrow$ that the two series are correlated

Delay and Integration Time

What delay?

Shift $\{y_i\}$ vs $\{x_i\}$ and calculate: r_k ; S_k ; C_k
…then look for the maximum confidence C_M
Time shift for C_M = delay between IFOs
Shift limits: ±10 ms (LLO-LHO light travel time)

$$r_k = \frac{\sum_i (x_i - \bar{x})(y_{i+k} - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2}\sqrt{\sum_i (y_{i+k} - \bar{y})^2}}$$

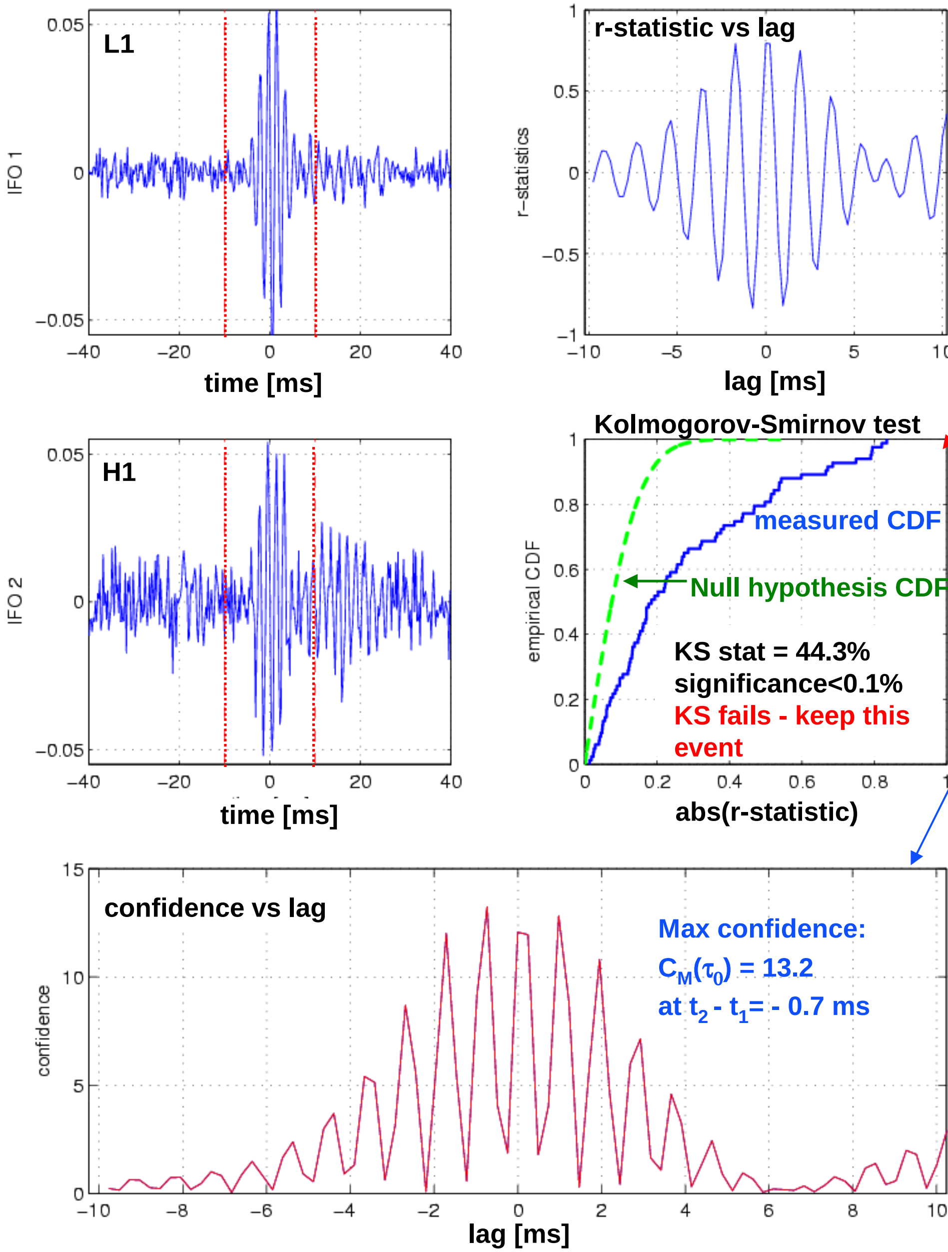
<u>Integration time τ:</u>

How long?

» If too small, we lose waveform information and the test becomes less reliable
» If too large, we wash out the waveform in the cross-correlation

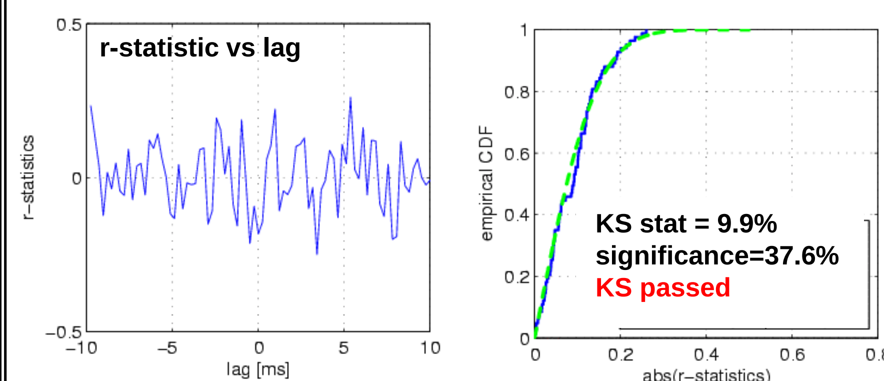
Test different τ's and do an OR of the results **(20ms, 50ms, 100ms)**

Simulated Sine-Gaussian Q=9, f_0=554Hz (passed through IFO response function)
h_{peak}=1e-18 [strain]; h_{rss}=5e-20 [strain/rtHz]

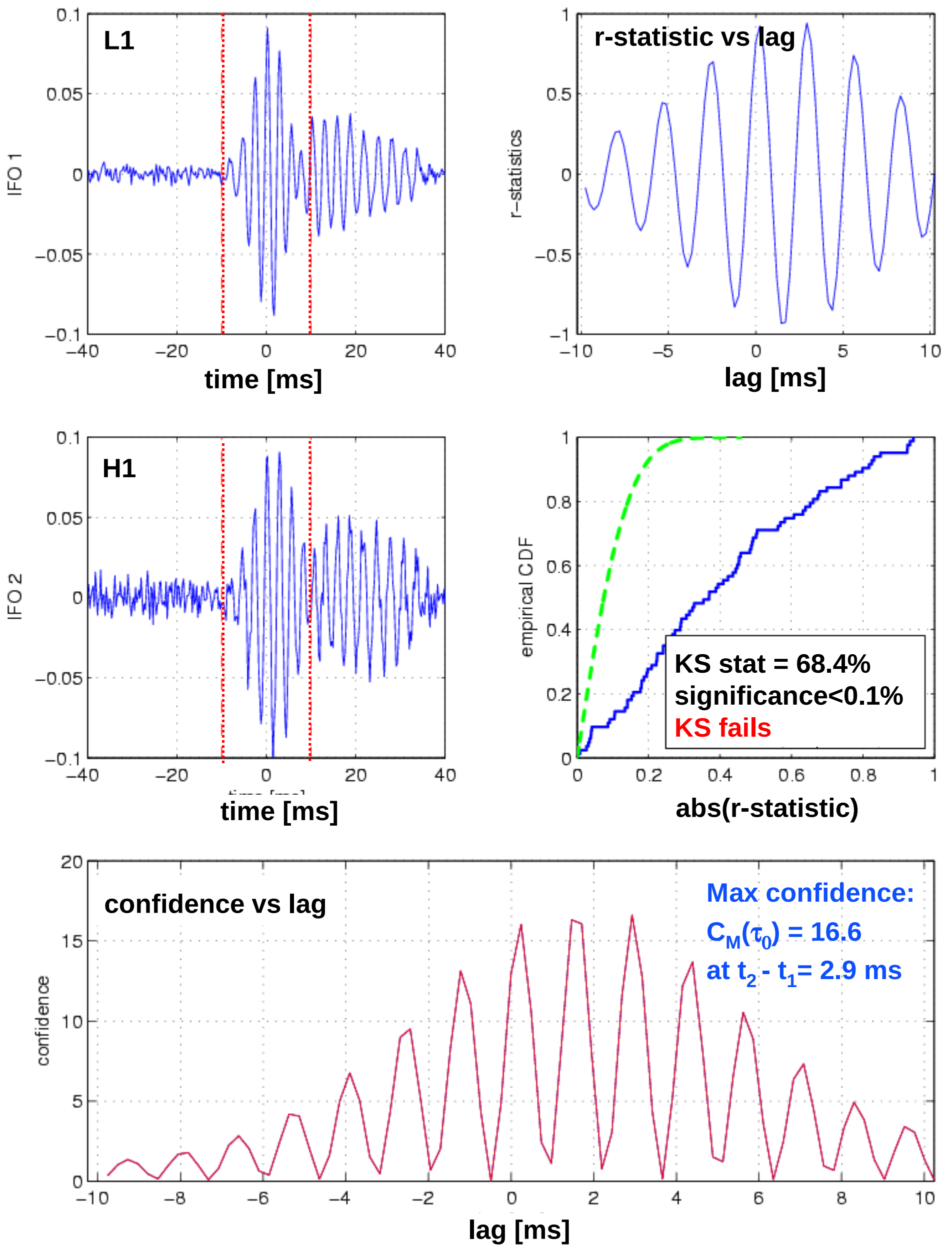


- L1-H1 pre-processed waveforms and r-statistic plot
  - integration time τ = 20 ms,
  - centered on the signal peak time.
- a Kolmogorov-Smirnov (KS) test states the $\{r_k\}$ distribution is NOT consistent with the null hypothesis.
  - » there is less than 0.1% probability that this distribution is due to uncorrelated series.
- On to the calculation of the confidence series and of its maximum C_M (j)

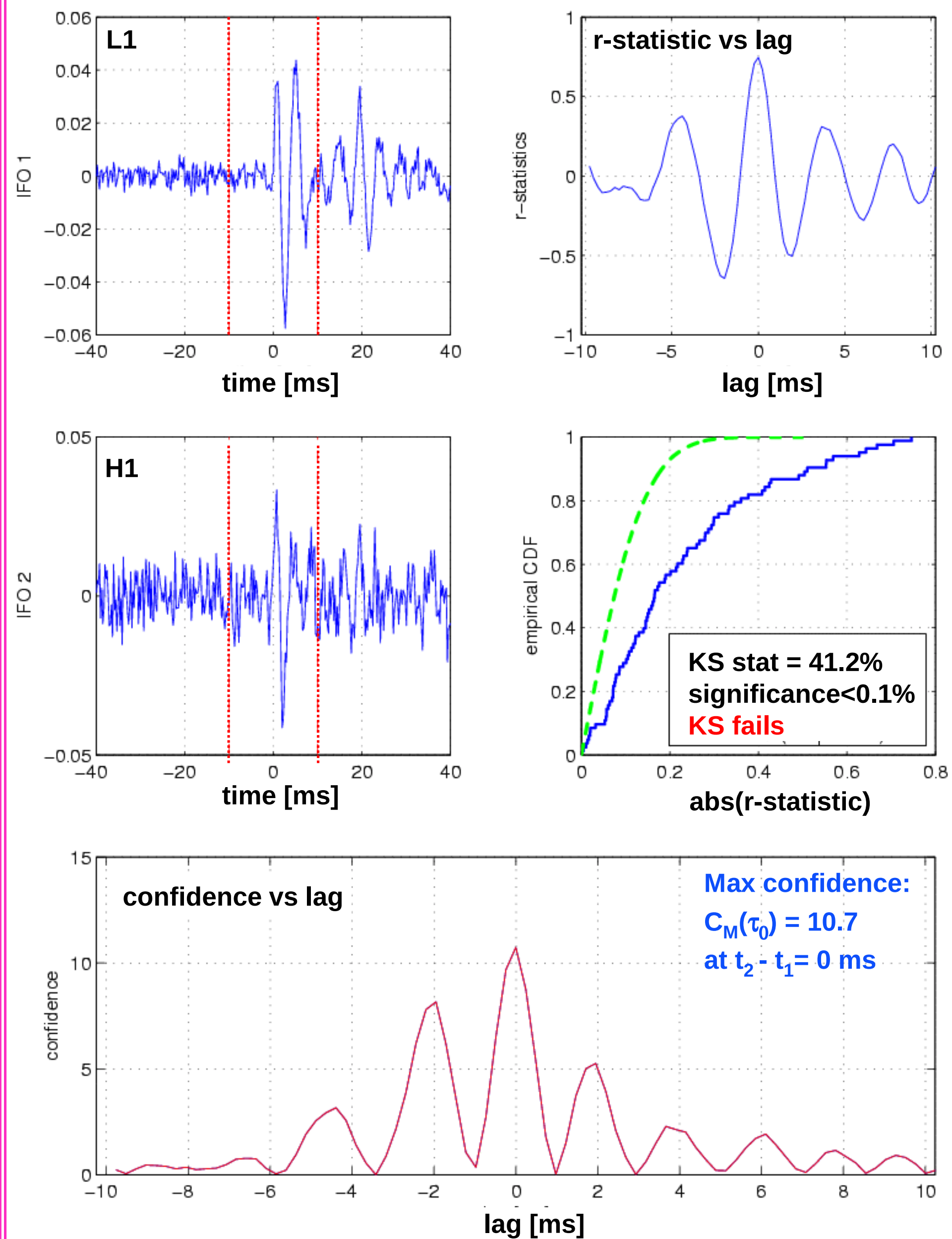
Noise only (no added signal)



Simulated Sine-Gaussian Q=9, f_0=361Hz
(passed through IFO response function) h_{peak}=1e-18

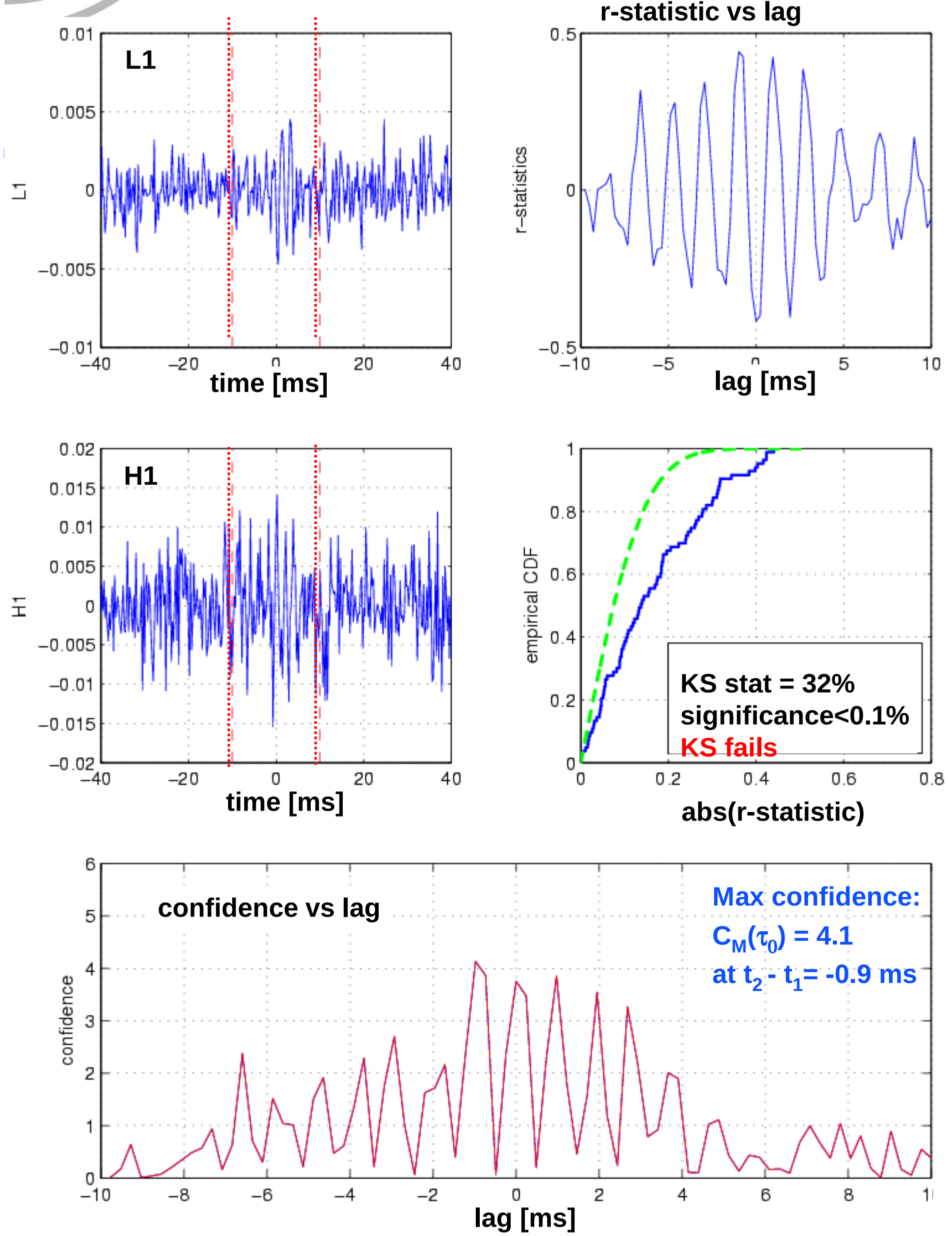


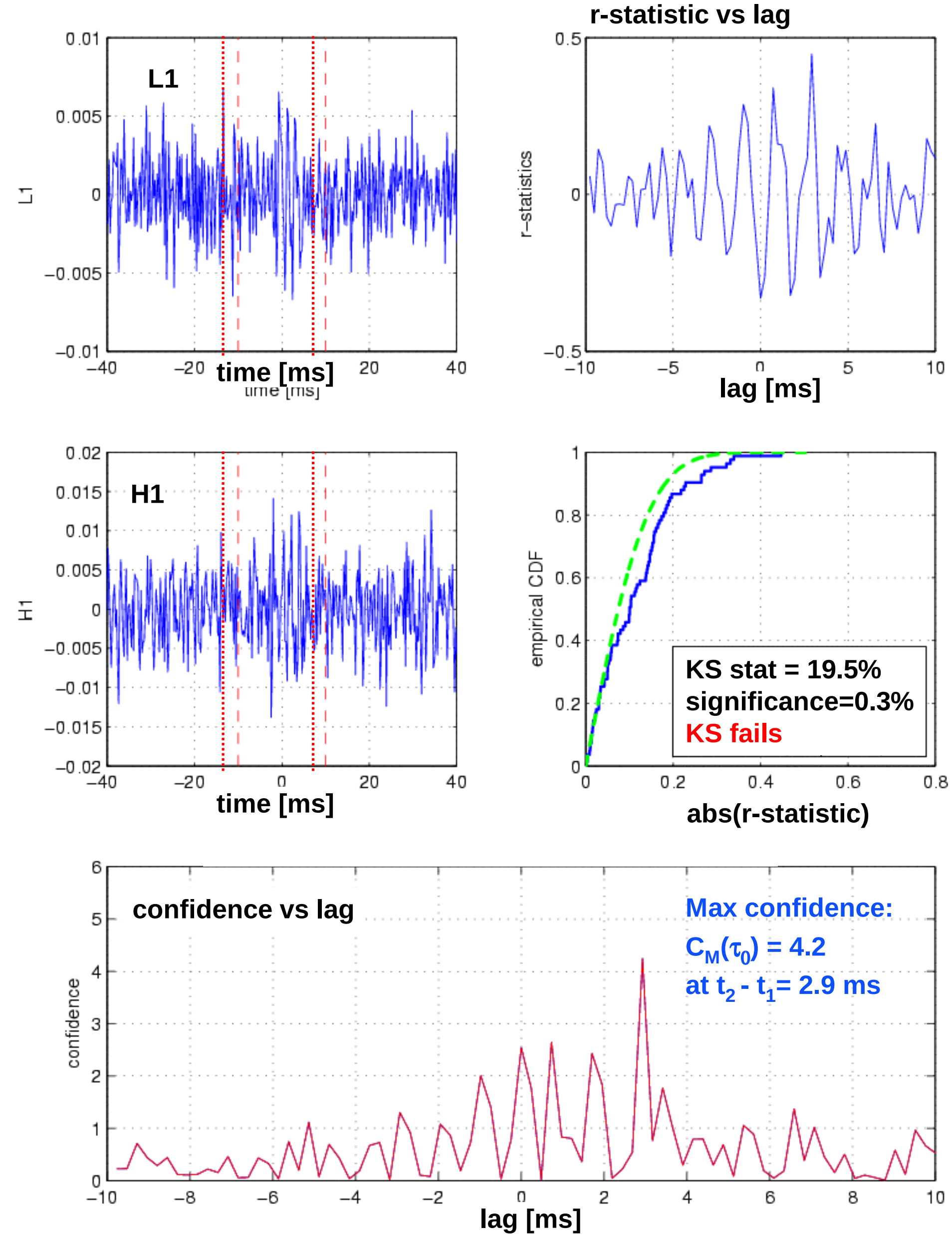
Simulated 1 ms Gaussian
(passed through IFO response function) h_{peak}=1e-18



Hardware injection Sine-Gaussian Q=9, f_0=554Hz - April 9 2003

h_{peak}~1e-18 h_{peak}~6e-21 [strain/rtHz]

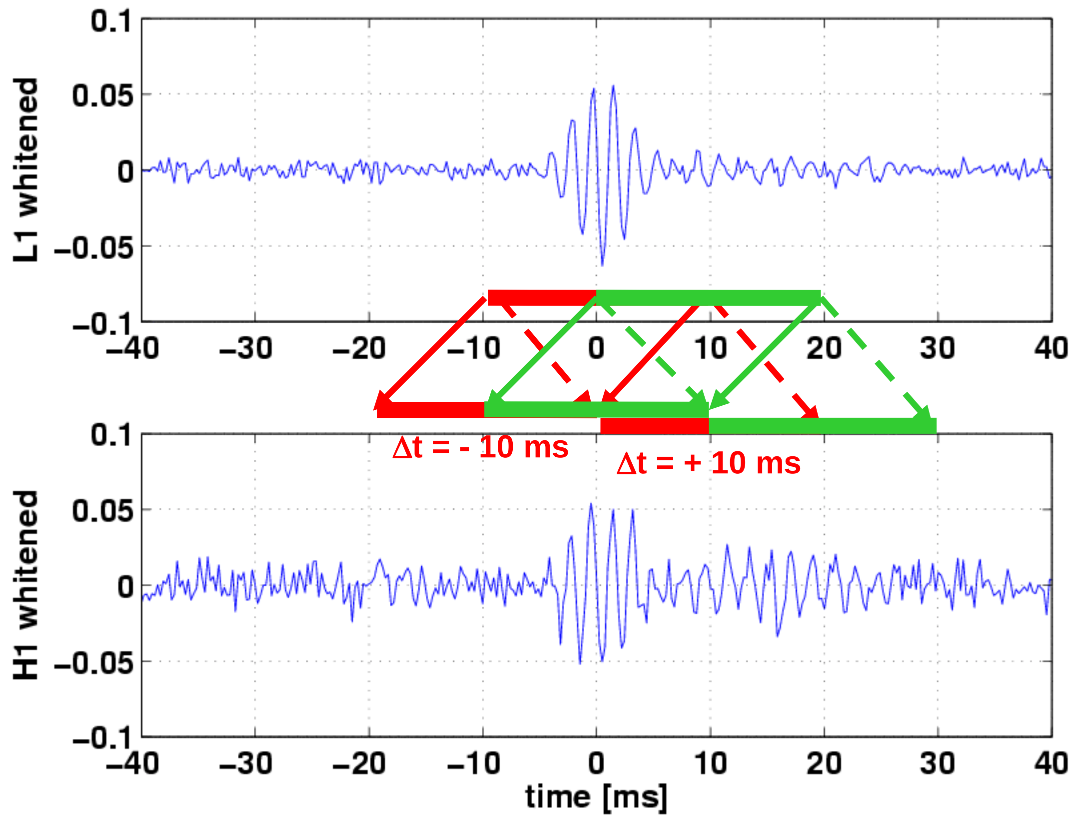




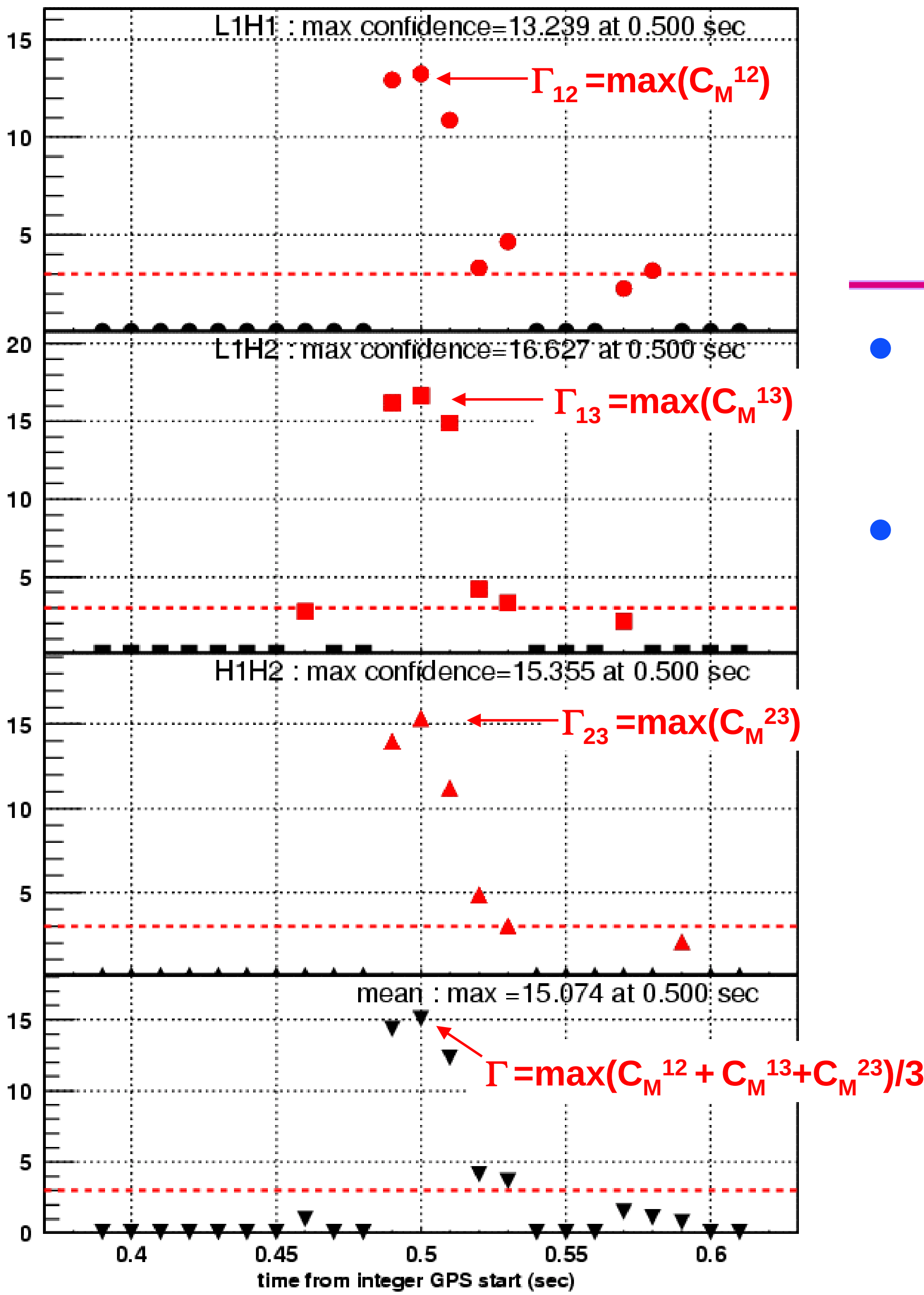
Scanning the Trigger Duration ΔT

» Partition trigger in $N_{sub}=(2\Delta T/\tau)+1$ subsets and calculate C_M (j) (j=1.. N_{sub})

» Use $\Gamma_{ab} = \max_j(C_M(j))$ as the correlation confidence for a pair of detectors over the whole event duration



$C_M(j)$ plots



- Each point: max confidence $C_M(j)$ for an interval τ wide (here: τ = 20ms)
- Define a cut (pattern recognition?):
  - 2 IFOs:
  - **Γ=max$_j$(C_M(j)) > β_2**
  - 3 IFOs:
  - **Γ=max$_j$(C_M^{12}+ C_M^{13}+ C_M^{31})/3 > β_3**

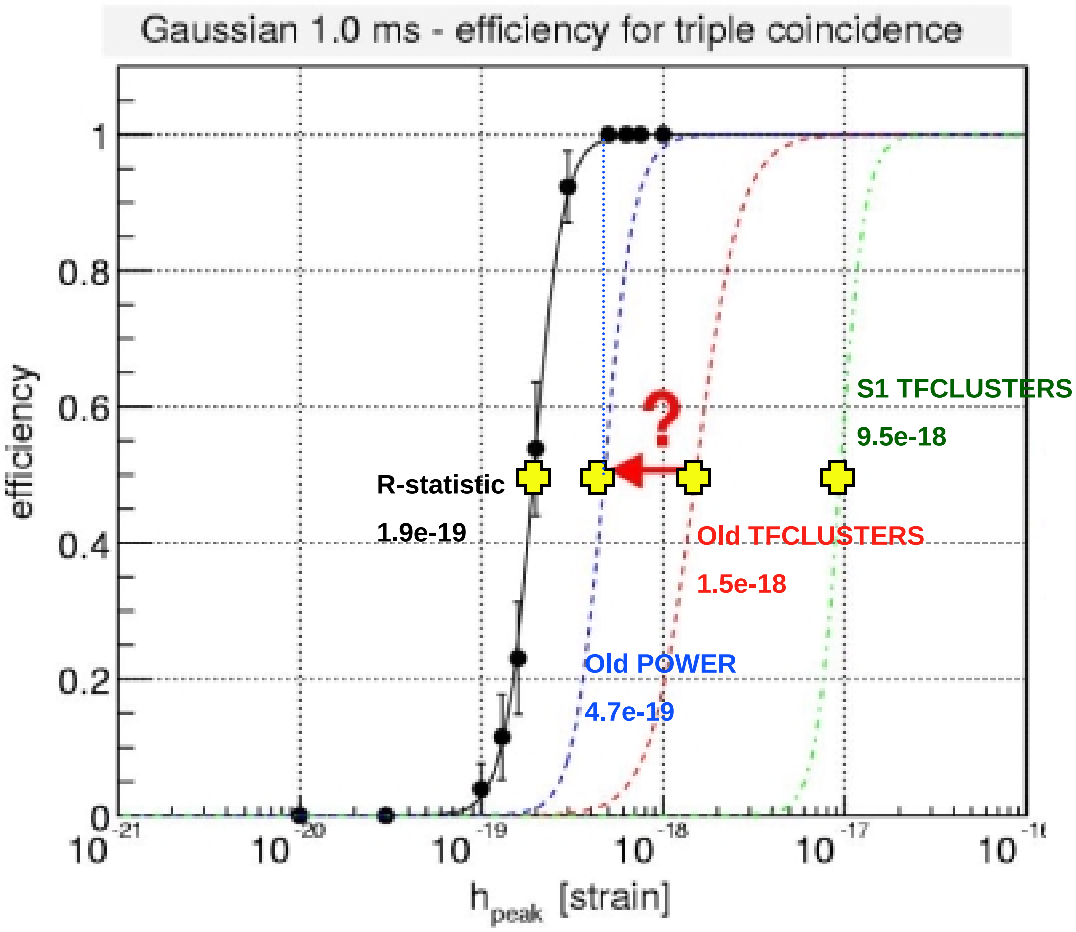
In general, we can have $\beta_2 \neq \beta_3$

β=3: 99.9% correlation probability

Open Issues

- Calibration
  - » Needed to account for waveform distortions due to frequency-dep calibration
- Time resolution
  - » Depends on waveform
  - » Affected by detector response function
    - – Implement phase calibration to match IFOs?
  - » Affected by pre-processing filters
    - – Do a second test pass with less aggressive filters?
  - » At the moment, no use is made of the delay time in assessing the confidence
    - – Implement a T-statistic test?

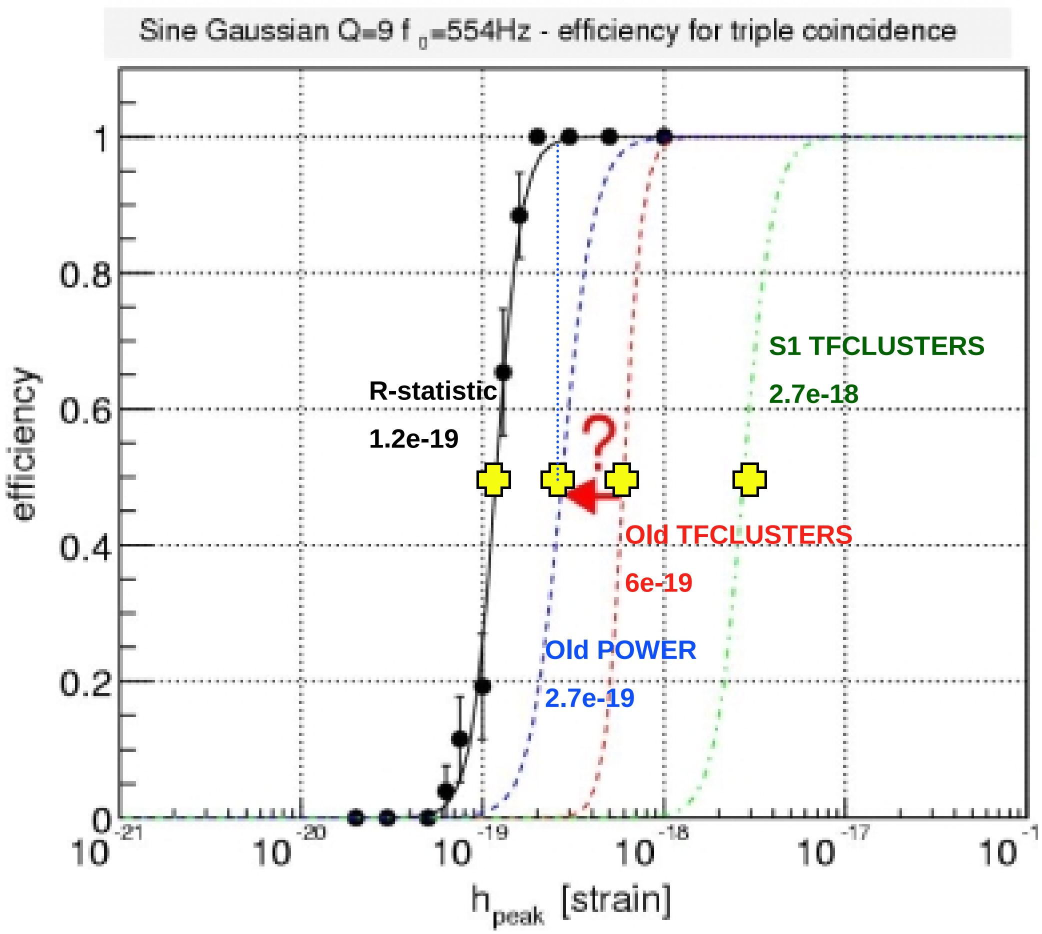
Triple-coincidence efficiency for 1 ms gaussians



- Same amplitude event injected in L1, H1, H2
- 0.2 sec around the injection time passed through the r-test.
- Sigmoid fit to efficiency curve:

  50% at h_{peak}=2e-19

Triple-coincidence efficiency for 554 Hz Q=9 sine gaussians

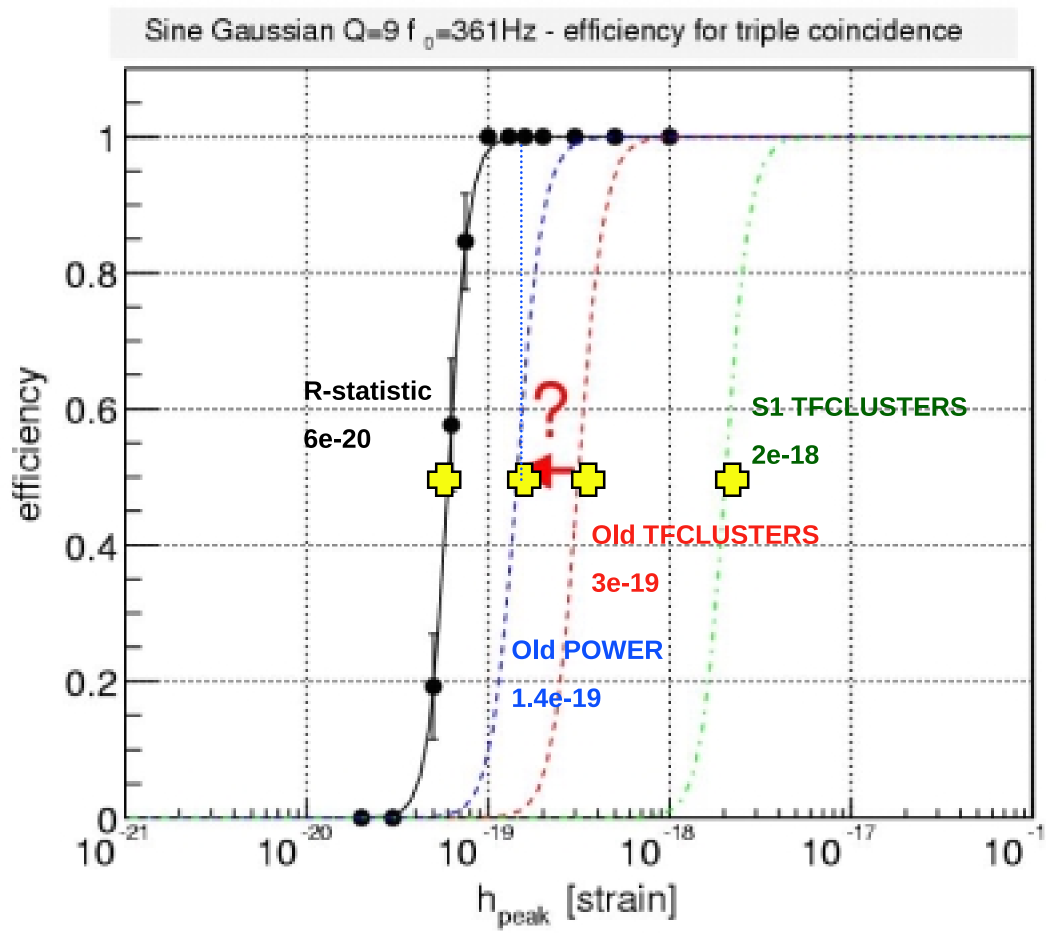


- Same amplitude event injected in L1, H1, H2
- 0.2 sec around the injection time passed through the r-test.
- Sigmoid fit to efficiency curve:

  50% at h_{peak}=1.2e-19

  h_{rss}=6e-21 strain/rtHz

Triple-coincidence efficiency for 361 Hz Q=9 sine gaussians



- Same amplitude event injected in L1, H1, H2
- 0.2 sec around the injection time passed through the r-test.
- Sigmoid fit to efficiency curve:

  50% at h_{peak}=6e-20

  h_{rss}=3.6e-21 strain/rtHz